УДК 004.02

ПОДХОД ДЛЯ ОБНАРУЖЕНИЯ СЕТЕВЫХ АНОМАЛИЙ

М.В. МАТЮШ
(Полоцкий государственный университет)

Статья посвящена решению задачи обнаружения аномальных ситуаций в процессе функционирования приложений. Показана актуальность решения этой задачи как меры повышения безопасности функционирования сетевых протоколов. Представлена классификация сетевых аномалий, ошибки реализации приложений, которые способствуют развитию вредоносного воздействия аномалии. Рассмотрена эффективность простых методов анализа сигнатур и протоколов передачи данных, показано их единство для решения задачи выявления аномалий. Предложен подход к выделению инвариантных признаков, свойственных нормальному функционированию сетевого протокола, а также подход реализации метода выявления аномалий в сообщении сетевого протокола как для обязательных полей, присущих всем сообщениям протокола, так и метод, четко детектирующий признаки в пакете сообщения, которые являются недопустимыми для анализируемого сетевого протокола.

Введение. Протокол передачи данных – набор соглашений логического уровня интерфейса, который определяет обмен данных между различными программами в компьютере. Эти соглашения задают единообразный способ передачи сообщений и обработки ошибок при взаимодействии программного обеспечения разнесённой в пространстве аппаратуры, соединённой тем или иным интерфейсом [2]. Сетевой протокол – набор правил, позволяющий осуществлять соединение и обмен данными между двумя и более включёнными в сеть устройствами. Разные протоколы зачастую описывают лишь разные стороны одного типа связи; взятые вместе, они образуют стек протоколов.

Проблемы с безопасностью клиентского приложения могут начаться еще до завершения его написания – в момент, когда выбираются протоколы и стандарты, по которым приложение будет работать. Часто проблемы связаны с попыткой разработчика сделать протокол слишком универсальным и описать в нем как можно больше функций. Из-за этого получается громоздкий многофункциональный протокол, реализовать который чрезвычайно сложно, еще сложнее соблюсти при этом безопасность в нем [2, 3].

Необходимо учитывать многие аспекты безопасности клиентских приложений: безопасность используемого протокола передачи данных, безопасность выбранного метода аутентификации и шифрования, качество кода клиентского приложения [2, 3]. Любое обращение клиентского приложения на сервер связано с некоторым обменом данных, в случае сетевой атаки на эти данные происходят сбои (аномалии), которые приводят к искажению передаваемых по сети данных с целью нарушения установившихся соединений или получения несанкционированного доступа к информационным ресурсам. В любом случае характер аномалий не предсказуем и может отражаться как в некорректной работе приложения, так и в потере и хищении ценной конфиденциальной информации.

На сегодняшний день не существует общего подхода к решению задачи обнаружения аномальных ситуаций в процессе функционирования приложений. Однако в условиях бурного развития информационных технологий и, как следствие, постоянной модернизации программного и аппаратного обеспечения решение частных задач обнаружения аномалий не может обеспечивать безопасность системы. Необходим более универсальный и вместе с тем научно обоснованный метод отслеживания состояний протоколов. Актуальность проблемы подчеркивается еще и тем, что даже в случае более или менее полного покрытия всей области наиболее важных проблем, связанных с обнаружением аномалий выбранного сетевого протокола, частными решениями, интеграция этих решений представляет собой трудноразрешимую комплексную задачу [4].

Классификация сетевых аномалий и методы их обнаружения. Известные сетевые аномалии настолько разнообразны, что единой классификации они не поддаются. Так, существует деление на активные и пассивные, внешние и внутренние, умышленные и неумышленные аномалии и т.д. Однако данные подходы не отражают всех характеристик изучаемого явления и являются ограниченными. Поэтому автором предлагается классификация сетевых аномалий с точки зрения объекта воздействия – информационной системы, включающей программно-аппаратный комплекс и сетевую инфраструктуру [5].

Согласно выбранному подходу можно разделить сетевые аномалии на две основные группы (рисунок).

К программно-аппаратным отклонениям относятся:

- аппаратные неисправности;
- ошибки конфигурирования;
- ошибки программного обеспечения;
- проблемы производительности оборудования.

Распространенные ошибки реализации сетевых протоколов в клиентских приложениях показаны в таблице [6].

Классификация сетевых аномалий

Распространенные ошибки реализации сетевых протоколов

Ошибка	Причина
Переполнение буфера, ошибки форматной строки, целочисленные переполнения	Часто возникают при разборе клиентом ответа на команду или даже при длинном приветственном сообщении сервера
Манипуляция данными	Возникает из-за подмены содержимого в браузерах и почтовых программах
Отказ в обслуживании	Возникает при разборе сложных или нестандартных данных и может привести к зависанию приложения. Чаще всего возникает из-за невозможности обработать исключительные или граничные ситуации
Недостаточная проверка данных	Например, недостаточная проверка пути доверия сертификата или обратный путь в каталогах, при котором серверное приложение может управлять тем, в какую папку попадет файл, загруженный клиентом
Утечка информации	Клиент передает данные о себе системе или пользователю. И передает больше, чем нужно

Нарушения сетевой безопасности включают в себя следующие аномалии:
- сканирование;
- атаки с целью отказа от обслуживания;
- вирусная активность;
- эксплуатация уязвимостей;
- анализаторы трафика;
- сетевые модификаторы.

Наибольший экономический ущерб операторам связи наносят атаки с целью перегрузки сетей или сервисов и сетевая вирусная активность.

При проверке сетевого трафика проверяют весь входящий и исходящий сетевой трафик. Ищется злонамеренный трафик, который может означать возможное нападение или неправомерное использование. Анализ сигнатур был первым методом, примененным для обнаружения вторжения. Он базируется на простом понятии совпадения последовательности с образцом. Во входящем пакете просматривается байт за байтом и сравнивается с сигнатурой (подписью) – характерной строкой программы, указывающей на характеристику вредного трафика. Такая подпись может содержать ключевую фразу или команду, которая связана с нападением. Если совпадение найдено, объявляется тревога. В противном случае в пакете отыскивается следующая подпись. Как только все подписи проверены, в память записывается следующий пакет, и процесс начинается снова. Каждый пакет сопровождается различными протоколами, которые могут разворачиваться и осматриваются согласно стандартам или RFC. Каждый протокол имеет несколько полей с ожидаемыми или нормальными значениями. Если что-нибудь нарушает эти стандарты, то вероятна злонамеренность. Если имеются нарушения протокола, например, если он содержит неожиданное значение в одном из полей, объявляется тревога. Анализ протокола использует детальное знание об ожидаемых или нормальных значениях в полях пакета для того, чтобы обнаружить вредоносный трафик. Анализ протокола очень отличается от анализа сигнатур, который использует известные характеристики атак для объявления тревоги [7].

Системы анализа сигнатуры имеют несколько важных сильных сторон. Во-первых, они очень быстры, так как полный анализ пакета – относительно тяжелая задача. Правила легко написать, понять и настроить. Кроме того, имеется поддержка компьютерного сообщества в быстром производстве сигнатур для новых опасностей. Эти системы превосходят все другие при поимке хакеров на первичном этапе. Простые атаки имеют привычку использовать некие предварительные действия, которые легко распознать.

В случае анализа протоколов имеем аналогичную ситуацию. Эта система тоже имеет свои положительные и отрицательные стороны. Из-за предпроцессов, требующих тщательной экспертизы протоколов, анализ протокола может быть довольно медленным. Кроме того, правила проверки для системы протокола трудно написать и понять. Можно даже сказать, что в этом случае приходится уповать на добросовестность производителя программы, так как правила относительно сложны и трудны для самостоятельной настройки. Более того, правила становятся все более и более сложными, часто игнорирующими общепринятые стандарты, протоколы и RFC.

На первый взгляд два метода обнаружения вторжения – анализ сигнатур и анализ протокола – кажутся весьма разными, но изучение проблемы показывает их некоторое сходство. В конце концов, эти инструменты безопасности исследуют форматированные данные об атаках и аномалиях.

Подход к выделению инвариантных признаков и выявлению аномалий сетевого протокола. Для анализа протокола необходимо классифицировать группы сообщений протокола так, чтобы выявить общие признаки сообщений, характерных только для конкретного протокола. На их основании строятся эталонные характеристики для протокола передачи данных.

Исследуется некоторое множество M всех возможных сообщений сетевого протокола передачи данных:

$$m_i \in M, i = \overline{1..I},$$

где m_i – любое сообщение сетевого протокола; I – количество всех возможных сообщений сетевого протокола.

Для любого сообщения m_i из этого множества можно определить априорное множество признаков Q_i, описывающих данное сообщение, в виде:

$$q_{ij} \in Q_i, j = \overline{1..J}.$$

Здесь q_{ij} – отдельный признак сообщения m_i сетевого протокола; J – априорное количество признаков, определяющих сообщение m_i.

Требуется по предъявленному набору значений признаков Q_i, т.е. описанию некоторого сообщения m_i из M, выделить общие признаки X этого сообщения, характерные для всего протокола.

Можно выделить следующие виды признаков, которые могут быть присущи всем сообщениям любого сетевого протокола, которые можно поделить на три непересекающихся подмножества признаков: обязательные X^F, дополнительные X^V и неиспользуемые X^W признаки. Представим их как

$$X = X^F \cup X^V \cup X^W. \quad (1)$$

Для выражения (1) справедливо выполнение условия непересечения множеств:

$$\begin{cases} X^F \cap X^V = \varnothing; \\ X^F \cap X^W = \varnothing; \\ X^V \cap X^W = \varnothing. \end{cases}$$

Данное разделение признаков X на три подмножества позволяет более детально определить стратегию проверки соответствия признаков.

Множество дополнительных признаков X^V содержит полный возможный набор признаков для сообщения сетевого протокола, а m_i может содержать только некоторые признаки из множества X^V или не содержать их вообще.

Помимо представления (1) признаки X можно поделить на целочисленные N, логические B и номинальные P:

$$X = N \cup B \cup P. \tag{2}$$

Так как каждое подмножество в (1), входящее в множество X описания сообщения сетевого протокола, может объединять в себе целочисленные, логические и номинальные признаки, соответственно, то оно может быть представлено выражением (2) и справедлива следующая система:

$$\begin{cases} X^F = N^F \cup B^F \cup P^F; \\ X^V = N^V \cup B^V \cup P^V; \\ X^W = N^W \cup B^W \cup P^W. \end{cases} \tag{3}$$

Система (3) расширяет выражение (2) и позволяет более детально произвести сравнение признаков по принадлежности к общему протоколу. Учитывая систему (3), опорная система признаков для анализа сетевого протокола принимает следующий вид:

$$X = N^F \cup N^V \cup N^W \cup B^F \cup B^V \cup B^W \cup P^F \cup P^V \cup P^W. \tag{4}$$

Объединив целочисленные, логические и номинальные признаки из выражения (4), можно получить систему:

$$\begin{cases} N = N^F \cup N^V \cup N^W; \\ B = B^F \cup B^V \cup B^W; \\ P = P^F \cup P^V \cup P^W. \end{cases} \tag{5}$$

Данное объединение возможно из-за того, что в совокупности (1) и (2) описывают один и тот же набор признаков с разных сторон их восприятия. Соответственно, данное объединение является результатом этих двух восприятий, что не противоречит факту того, что любой из признаков в выражении (2) может состоять из обязательных, необязательных и неиспользуемых признаков сетевого сообщения.

Система (5) описывает достаточный набор признаков для детектирования сообщения сетевого протокола. Ниже будут определены общие логические признаки для всех сообщений диалекта с дальнейшими трансформациями системы (5) с целью уменьшения избыточности требуемых данных при детектировании сетевого протокола.

В соответствии с подходом анализа сетевого протокола к логическим признакам B_i сообщения m_i относится представление полей сообщений сетевого протокола.

Покажем эти признаки в трехмерной логике: «истина» – в 1, «ложь» – в 0, а «неопределенность» – в минус 1. Логические параметры b_j из множества B_i в трехмерной логике принимают значения:

$$b_j = \begin{cases} 1, \text{если } b_j \in B^F; \\ 0, \text{если } b_j \in B^W; \\ -1, \text{если } b_j \in B^V, \end{cases} \tag{6}$$

где j – определяет номер поля в сообщении сетевого протокола.

Логические признаки B в состоянии качественно описать сетевой протокол. Это обусловлено тем, что поля сообщения сетевого протокола несут описание как самого протокола, так и его возможных версий, как правило, отличающихся набором полей. Из-за наличия дополнительных признаков наполнение определенных полей варьируется.

Нижеприведенный метод синтеза можно также применить и для целочисленных, и номинальных признаков сообщения, но мы выбрали именно логические как наиболее информативные сущности, описывающие сетевой протокол.

Для перевода из трехзначной в двухзначную логику представим логический признак B в виде строгого ${}^{f}B$ и расширенного ${}^{e}B$ представления признаков, тогда из (6) элементы множества ${}^{f}B$ принимают значения:

$$ {}^{f}b_j = \begin{cases} 1, \text{ если } {}^{f}b_j \in B_z^F; \\ 0, \text{ если } {}^{f}b_j \in B^V \cup B^W, \end{cases} \tag{7} $$

а, соответственно, элементы множества ${}^{e}B$ принимают значения:

$$ {}^{e}b_j = \begin{cases} 1, \text{ если } {}^{e}b_j \in B^F \cup B^V; \\ 0, \text{ если } {}^{e}b_j \in B^W. \end{cases} \tag{8} $$

Согласно выражениям (7) и (8) можно характеризовать сетевой протокол по строгой и расширенной выборке логических признаков.

Для получения описания сетевого протокола объединим все логические признаки всех сообщений. Учитывая то, что логические признаки могут иметь строгое и расширенное представления, показанные в (7) и (8), можем получить логические признаки, характерные для сетевого протокола, и представить выражение в виде:

$$ B = \bigcap_{y=1}^{Y} {}^{f}B_y, \tag{9} $$

или

$$ B = \bigcup_{y=1}^{Y} {}^{e}B_y, \tag{10} $$

где Y – количество возможных сообщений выбранного сетевого протокола.

Выражения (9) и (10) определяют логический строгий и логический расширенный наборы признаков для описания сетевого протокола. Далее предложим реализацию метода для выявления аномалий сетевого протокола сопоставлением логических признаков, характерных для всех его сообщений.

Пусть функция $L(m_i)$ возвращает логическое описание сообщения сетевого протокола (наличия полей). Теперь попытаемся выявить аномалию по сопоставлению строго логического описания сетевого протокола:

$$ L(m_i) = {}^{f}B \cap L(m_i). \tag{11} $$

Если верно выражение (11), то полученный/отправленный пакет не содержит аномалий и является стандартным решением для выявления аномалий. Учет же дополнительных полей сообщений сетевого протокола позволяет дать более точные результаты поиска аномалий сетевого протокола по сравнению с выражением (11).

Определим условие соответствия сообщения m_i диалекту по расширенному описанию сетевого протокола:

$$ L(m_i) = {}^{e}B \cup L(m_i). \tag{12} $$

Выражение (12) также можно интерпретировать как сравнение с неиспользуемыми полями $\overline{{}^{e}B}$ сообщения, учитывающее логические признаки, которые однозначно не относятся к соответствующим признакам сетевого протокола.

Заключение

1. Показана актуальность решения проблемы возникновения аномалий, возникающих в сетевых протоколах, связанных с посторонним воздействием на них с целью проведения дестабилизации работы клиентского приложения или хищения конфиденциальной информации.

2. Представлена классификация сетевых аномалий, ошибки программистов, которые способствуют появлению аномалий, а также возможные варианты обнаружения аномалий для протоколов передачи данных.

3. Предложен подход к выделению инвариантов, которые могут описывать сетевой протокол на основании выбора общих признаков для всех его сообщений, и объединению этих признаков в строгие и расширенные характеристики этого сетевого протокола.

4. Предложена реализация метода выявления аномалий при анализе пакетов сетевого протокола путем сопоставления логических признаков, учитывающих обязательные и дополнительные логические признаки. Предложен учет дополнительных признаков, которые являются необязательными, но которые позволяют четко определить признаки, чуждые анализируемому сетевому протоколу. Наличие данных признаков свидетельствует об аномалиях в работе сетевого протокола.

ЛИТЕРАТУРА

1. Протокол передачи данных // Википедия: свободная энциклопедия [Электронный ресурс]. – 2002. – Режим доступа: http://ru.wikipedia.org/wiki/. – Дата доступа: 30.09.2010.
2. Безопасность сетевых протоколов // Спецвыпуск: Хакер, номер #058, с. 058-084-1 [Электронный ресурс]. – 1999. – Режим доступа: http://www.xakep.ru/magazine/xs/058/084/1.asp. – Дата доступа: 30.09.2010.
3. Безопасность распространенных прикладных сетевых протоколов: взгляд со стороны клиента // Информационная безопасность [Электронный ресурс]. – 2000. – Режим доступа: http://securityvulns.ru/ /articles/xs/network_client_sec.asp. – Дата доступа: 30.09.2010.
4. Баранов, П.А. Обнаружение аномалий на основе анализа однородности параметров компьютерных систем: дис. ... канд. техн. наук: 05.13.19 / П.А. Баранов. – СПб., 2007. – 155 с. РГБ ОД, 61:07-5/2850клиента // Научная электронная библиотека [Электронный ресурс]. – 2003. – Режим доступа: http://www.lib.ua-ru.net/diss/cont/171894.html. – Дата доступа: 30.09.2010.
5. #283 Сетевые аномалии // ОО «НАГ» [Электронный ресурс]. – 2000. – Режим доступа: http://www.nag.ru/ /articles/reviews/15588/setevye-anomalii.html. – Дата доступа: 30.09.2010.
6. Сводка уязвимостей в безопасности // Информационная безопасность [Электронный ресурс]. – 2000. – Режим доступа: http://www.security.nnov.ru/. – Дата доступа: 30.09.2010.
7. Анализ сигнатур или анализ протоколов, что лучше? // Портал компьютерных статей [Электронный ресурс]. – 2000. – Режим доступа: http://computerlib.narod.ru/html/ids.htm. – Дата доступа: 30.09.2010.

APPROACH FOR ANOMALY DETECTION

M. MATSIUSH

The article is devoted to solving the problem of detection of anomaly situations in the operation of applications. It shows the urgency of solving this problem as measures to enhance safe operation of network protocols. It's represented the classification of network anomalies, errors, implementation of applications that promote harmful effects of the anomaly. It's considered the efficiency of simple methods of signature analysis and data transmission protocols and showed their unity in solving the problem of anomaly detection. It's suggested an approach to the selection of invariant features characteristic of the normal functioning of the network protocol. It's suggested an approach to the implementation of the method of anomaly detection in a communication network protocol as for mandatory fields are common to all protocol messages and the method clearly detecting signs of the package the messages that are not valid for the analyzed network protocol.